NAME

rdseed − Read an FDSN SEED format volume

SYNOPSIS

rdseed -{aclsSRtdoxzei} [-f filename] [v (n)]

OPTIONS

The following options may be selected:

−a retrieve the abbreviation dictionaries.

−b set the seismogram ouput buffer size.

−c retrieve volume table of contents.

−l list contents of each record in the volume.

−q set the rdseed output directory

−Q select the data quality flag

−s retrieve all station and channel header information.

−S extract a "weed" compatible station information file.

−R retrieve all response information in a "evalresp" compatible format.

−t construct a list of all event start/stop times and starting record numbers.

−d dump the data records

−o data output format 1=SAC, 2=AH, 3=CSS, 4=Mini-SEED, 5=SEED, 6=SAC ASCII, 7=SEGY

−x read the summary file

−z check reversals. Argument is btw 1..3

−e extract event data (blockette 71) from seed volume.

−i ignore network codes.

−v specify the volume number [1-n] to process (default is 1).

−f specify the input filename (default is /dev/nrst0).

−g specify alternate response file(s), see Alternate Response Files below. This takes precedence over any files indicated with the ALT_RESPONSE_FILE environment variable.

DESCRIPTION

rdseed reads from the input tape (or other input file) in the format defined by the Federation of Digital Seismographic Networks (FDSN), popularly known as the Standard for Exchange of Earthquake Data (SEED). According to the command line function option specified by the user, **rdseed** will read the volume and recover the volume table of contents (-c option), the set of abbreviation dictionaries (-a option), or station and channel information and instrument response tables (-s option).

The desired volume on tape (or in a file) can be selected with the (-v option). The first volume is "1" , the second "2" and so forth. (Note: this option is only used when a physical volume contains more than one logical volume.

Two additional options allow access to detailed information concerning the actual contents of the volume (rather than reading files purporting to contain such information from the volume, for which one uses the -c option). The first of these options (-t) writes out a list of data start and stop times along with the starting records at which those data may be found. (For the purposes of this discussion, a **record** is defined as a set of station/channel/time continuous data records.) The other option (-l) is primarily a diagnostic tool; it writes a description of every record on the volume.

In order to extract data from the SEED volume for analysis by other packages, the user must run **rdseed** in User Prompt Mode which is discussed below. As data is extracted from the SEED volume, **rdseed** looks at

the orientation and sensitivity of each channel and determines if channel polarity is reversed. Refer to the description of blockettes 52 and 58 in the SEED manual for a description of reversed polarity. A negative sensitivity in blockette 58 is indicative of a reversed polarity. The user can request that reversed channels be corrected. This correction is a simple multiplication by -1.0 . The output response file information is NOT adjusted for channels which rdseed inverts.

Output Header Corrections
rdseed does correct the header information in the SAC, AH and CSS files as follows:
SAC headers contain both scale factor and Dip/Azimuth information. If a Dip/Azimuth reversal has been corrected, RDSEED will correct the Dip/Azimuth information in the SAC header. If a negative sensitivity is corrected, the header SCALE value is also made positive. If RDSEED finds both types of reversals, NO corrections to the data or header are made.

CSS headers are handled in the same way as SAC headers. The header information is found in CSS formatted relation files that are seperate from the data files. These files are named RDSEED.WFDISC, RDSEED.SITE and RDSEED.SITECHAN.

The AH format presents a problem. There is no place for Dip/Azimuth information. Headers contain only a scale factor. Therefore when a Dip/Azimuth reversal occurs, there is no place in the header to indicate that occurrance. Therfore the following rules are followed. If the user has not requested dip/azimuth checking and one is detected, **rdseed** will invert the sign of the scale factor in the AH header. If the user has asked for inversion checking, the data will be inverted and NO change in the scale factor will occur. Negative sensitivity is handled the same way as in the SAC and CSS headers. Again, if **rdseed** finds both types of reversals, NO corrections will be made.

For SAC and AH files, units of response are in Counts/M

User Prompt Mode
When rdseed is run without any options, the user is prompted for all of the options as well as some additional options.

Input Device (/dev/nrst0) :
the input device can be changed from the default of /dev/nrst0. Note that when a taoe is being accessed, it is best to use the norewind device. This allows rdseed to search for multiple volumes on tape accross tape file marks.

Output Device (stdout) :
the output device can be changed from the default of stdout.

Volume # [(1)-N] :
volume number can be changed from the default value of 1.

Options [acsSrRtde] :
one of the options must be selected. All except "d" and "e" have the same meaning as the corresponding command line options.

the "e" option will generate a "weed" compatible event file.

If the "d", "s" or "S" option is selected, further selection can be made.

Summary File (None) :
a rdseed summary file can be selected as input for rdseed. Used in conjunction with "weed" to specify station time windows.

Station List (ALL) :
a list of selected stations separated by spaces or commas. Wildcard substitution of characters "*" , "?" and "." is performed. A station name can be that of an alias whose name is defined in a file

whose filename is specified by the enviornment variable SEEDALIAS. See details below.

Channel List (ALL) :
a list of selected channels separated by spaces or commas. Wildcard substitution of characters "*" ,"?" and "." is performed.

Network Code (ALL) :
a list of selected networks separated by spaces or commas. Enter ?? to disable network processing.

Location Code (ALL) :
a list of selected location codes separated by spaces or commas. Enter CR to disable location code processing. If the "d" option is selected, further selection can be made.

Output Format [(1=SAC),2=AH,3=CSS, 4=miniseed, 5=seed,
6=SAC ascii, 7=SEGY] : select output format as SAC, AH, CSS 3.0 binary, mini seed, seed, SAC ascii or SEGY; default is SAC.

Check Reversal [(0=No), 1=Dip.Azimuth, 2=Gain, 3=Both]:
select signal reversal checking; default is No.

Select Data Type [(E=Everything), D=Data of Undetermined State, R=Raw waveform Data, Q=QC'd data] :
select data quality to processes.

Start Time(s) (FIRST) :
a list of seismogram start times of the form YYYY.DDD.HH:MM:SS.FFFF or YYYY/MM/DD.HH:MM:SS.FFFF separated by spaces. YYYY may be YY i.e. "90" for "1990". Least significant parts may be omitted, in which case they become zero i.e. "90.270" is time 00:00:00.000 of the 270th day of 1990.

End Time(s) (LAST) :
a list of seismogram end times of the same form as start times. Each start time (except the last one) must have a corresponding end time. If the last start time does not have a corresponding end time, the end time is assumed to be the last time in the volume.

Sample Buffer Length [2000000]:
each seismogram is assembled in a sample buffer prior to output. The size of the buffer can be changed. This number is the number of samples. If the length is too small for the any of the requested seismograms an error message will inform the user that the waveform is being broken into pieces. The user may increase the size of the buffer to avoid this problem.

Extract Responses [Y/(N)] :
channel response extraction is selected via this prompt.

Alternate Response Files

The user can specify that station configuration and responses be taken from another SEED volume. This will allow for two possible scenarios to be accomplished. First, if station information is found to be in error, the user can use the headers from another SEED volume that is known to be correct. Second, if the SEED volume is a "data-only" volume (MINI-SEED), the user can use a "data-less" SEED volume and the "data-only" SEED volume to extract the desired waveforms. So you say, "how do I do this?". The environment variable ALT_RESPONSE_FILE should be defined with its value being the path name of the SEED volume from which responses should be taken.

The Alias File

An alias file can be defined which contains a list of station alias names. The first word in each line of the file is the alias. The following words are station names which will match the corresponding alias. The alias file name must be defined in the SEEDALIAS environment variable. For example, the file rdseed.alias contains the following:

CHINA BJI XIAN SHNG

All references to station CHINA will match station BJI, XIAN or SHNG.

The source code is written in such a way that it will run, with recompilation, on both big-endian (e.g., Sun) and little-endian machines (e.g., VAX) under UNIX. This program has not yet been tested under VAX VMS.

Time Tear Tolerance

Normally, the tolerance for determining time tears is found in the station header information. Some stations may have clocks that wander excessively. This may cause time tears in the data. The drift tolerance can be adjusted by defining an environment variable called SEEDTOLERANCE. Its value is multipled by a drift tolerance that is contained in the SEED volume for each channel. Thus a value of 3.0 will increase the drift tolerance by a factor of three. Drift is defined in units of "seconds per sample" and is typically around .00005 sec/smp.

Rdseed Alert message file

When rdseed determines that data reversal is necessary and the user specifies that rdseed should reverse the data, rdseed creates a file with the data reversal information inside. This information includes the file name where the reversal was applied. When you exit the program, a message is displayed reminding you to look at this file.

This file is called rdseed.alert.log and is located in the startup directory.

Rdseed error logging

All rdseed error messages are logged to a file, called rdseed.error.log with the date. This file is appended to throughout the day.

Data Recovery

There are two necessary steps to recovering seismograms from a SEED tape.

The first step consists of finding out what is on the tape; one does this by using any of the command line options -c or -t, to list the station and channel names, starting times, and record numbers of the seismograms contained in the volume.

Seismic data are recovered from SEED tapes in the second step. Using the station, channel and time information, use User Prompt Mode to select start and stop times for individual seismograms.

Seismogram files are written to the current directory with names of the form

yyyy.ddd.hh.mm.ss.ffff.SSSSS.CCC.SAC for SAC Files
yyyy.ddd.hh.mm.ss.ffff.SSSSS.CCC.AH for AH Files
yy.ddd.hh.mm.ss.ffff/SSSSS.CCC.w for CSS Files
mini.seed for mini seed
yyyy.ddd.hh.mm.ss.ffff.SSSSS.CCC.SAC_ASC for sac ascii
yyyy.ddd.hh.mm.ss.ffff.SSSSS.CCC.SEGY for SEGY

where yyyy is the year, ddd is the Julian day, hh.mm.ss.ffff is the time of day of the start of the first record, SSSSS is the station name, and CCC is the component name for the particular seismogram being recovered. This seismogram file naming convention was chosen to provide unique names to output files without user intervention; however, the large number of files which can be generated to a single directory might cause problems for some operating systems. Notice that CSS uses a slightly different format that puts channel data in subdirectories. This is due to a limitation in the filename field in the CSS database. For CSS there

are three additional files created, RDSEED.WFDISC, RDSEED.SITE and RDSEED.SITECHAN. These files are used by the CSS database system.

Be aware that rdseed always appends onto the mini.seed file. You need to manually remove this file if you wish to start over.

Recovering auxiliary data from an FDSN SEED volume.

One may also retrieve the set of abbreviation dictionaries or the set of station information tables from an FDSN SEED volume. The former is accomplished with the command
rdseed -af inputfile
while the latter results from
rdseed -sf inputfile
Output from these commands should be redirected to files, as in examples 6 and 7.

DIAGNOSTICS

Various warnings and error messages are issued to the standard error device by the procedure. Typical response of the procedure to a warning condition is to write a message to the standard error device and then to continue execution. An error condition, on the other hand, will cause a message to be generated to the standard error device followed by immediate termination of the procedure.

EXAMPLES

1. Reading the table of contents from a volume.

```
rdseed -cf /dev/rmt8 > tape.contents &
        or
rdseed
Input Device (/dev/rst0)      : /dev/rmt8
Output Device (stdout)        : tape.contents
Volume #        [(1)-N]     :
Options         [acsSrRtde]  : c
```

reads the table of contents from the tape on device /dev/rmt8 into a file called tape.contents; the job is run in the background (command line example). Output directed to "stderr", such as error messages, comes to the terminal. Table-of-contents entries have the format

```
 AFI   LPZ     1988,228,00:00:00.68         1        147
```

where the columns represent station name, channel name, start time, location flag (here, a blank), subsequence number, and starting record number.

2. Determining event start/stop times on a volume.

```
rdseed -tvf 2 image > image.times &
        or              rdseed
Input Device (/dev/rst0)      : image
Output Device (stdout)        : image.times
Volume #        [(1)-N]     : 2
Options         [acsSrRtde]  : t
```

reads a disk file called "image" and creates a table containing starting record numbers, station and channel names, start and stop times of events, nominal sample rate, calculated sample rate and numbers of samples for that file. Output is written to the file "image.times". The lines of the output table have the format

```
119  ANMO  LPZ  1988,230,00:00:00.6800  1988,230,18:35:04.6800  1.00  66904
```

where the columns represent station name, channel name, start time, end time, sample rate in samples per second, number of samples, and starting record number for these data.

3. Creating a detailed list of the contents of a volume.

```
rdseed -lf /dev/rmt11 > tape.list &
```

reads a tape on drive mt11 and writes a list of the contents of each record to a file called tape.list. The job

is run in the background.

4. Reading all data from a tape.

```
rdseed
Input Device (/dev/rst0)       : /dev/rmt8
Output Device (stdout)         : tape.extraction.list
Volume #         [(1)-N]       :
Options          [acsSrRtde]   : d
Station List (ALL)             :
Channel List (ALL)             :
Output format [(1-SAC),2-AH...]:
Start Time(s) (FIRST)          :
End Time(s) (LAST)             :
Sample Buffer Length [2000000] :
Extract Responses [Y/(N)]      :
```

reads all seismograms from the tape on device /dev/rmt8 into the current directory, writes "stdout" output to a file.

5. Reading the abbreviation dictionaries.

```
rdseed -af tape.image > tape.abbreviation.dictionaries
        or              rdseed
Input Device (/dev/rst0)      : tape.image
Output Device (stdout)        : tape.abbreviation.dictionaries
Volume #         [(1)-N]      : 1
Options          [acsSrRtde]  : a
```

extracts the abbreviation dictionaries from a tape image stored on disk, sends the result to a file, and runs in the foreground.

6. Reading station information.

```
rdseed -sf /dev/rmt8 > tape.station.information &
        or              rdseed
Input Device (/dev/rst0)      : /dev/rmt8
Output Device (stdout)        : tape.station.information
Volume #         [(1)-N]      : 1
Options          [acsSrRtde]  : s
```

recovers station and channel location and response information from the tape on device /dev/rmt8, writes the information to a file, and in the command line form runs in the background.

7. Reading specific station/channel/time information.

```
rdseed
Input Device (/dev/rst0)      : /dev/rmt8
Output Device (stdout)        : tape.station.information
Volume #         [(1)-N]      : 2
Options          [acsSrRtde]  : d
Summary File (None)               :
Station List (ALL)            : BJI YKW1
Channel List (ALL)            : *Z
Network List (ALL)                :
Output format [(1-SAC),2-AH..]:
Start Time(s) (FIRST)         : 1990,270,20:30
End Time(s) (LAST)            : 1991/2/1
```

```
Sample Buffer Length [2000000]: 3000000
Extract Responses [Y/(N)]     : Y
```

reads all seismograms from stations BJI and YKW1, all Z channels, from year 1900, julian day 270, hour 20, minute 30 to Febuaury 1, 1991. The buffer size was increased to 3 million samples and the channel response information will be output.

FILES

yyyy,ddd,hh.mm.ss.ffff.SSSSS.CCC - seismogram files as described above

mini.seed - file of mini seed data

rdseed.err_log.MM.DD.YY - a log of all error messages created during the day's usage.

rdseed.alert_log - file created when rdseed reverses the data.

SEE ALSO

Halbert, S. E., R. Buland, and C. R. Hutt (1988). Standard for the Exchange of Earthquake Data (SEED), Version V2.0, February 25, 1988. United States Geological Survey, Albuquerque Seismological Laboratory, Building 10002, Kirtland Air Force Base East, Albuquerque, New Mexico 87115. 82 pp.

O'Neill, D. (1987). IRIS Interim Data Distribution Format (SAC ASCII), Version 1.0 (12 November 1987). Incorporated Research Institutions for Seismology, 1616 North Fort Myer Drive, Suite 1440, Arlington, Virginia 22209. 11 pp.

Tull, J. (1987). SAC User's Manual, Version 10.2, October 7, 1987. Lawrence Livermore National Laboratory, L-205, Livermore, California 94550. ??? pp.

BUGS

The -d option dumps all records .

Address bug reports to Chris Laughbon at IRIS DMC, 206-547-0393, or send ARPAnet mail to chris@iris.washington.edu . Please DO NOT "fix bugs" locally without telling the author about them. **Please DO NOT redistribute this program.** Address requests for the program to IRIS at the numbers given above.

AUTHOR

Dennis O'Neill; Allen Nance, IRIS